

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Thurber, Denise
Zeng, Bin
- (ii) TITLE OF INVENTION: Coding Sequences of the Human
BRCA1 Gene
- (iii) NUMBER OF SEQUENCES: 72
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Howrey & Simon
(B) STREET: 1299 Pennsylvania Avenue N. W.
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: USA
(F) ZIP: 20004
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Albert P. Halluin
(B) REGISTRATION NUMBER: 25,227
(C) REFERENCE/DOCKET NUMBER: 5371.34.US00
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-463-8109
(B) TELEFAX: 650-463-8400

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5711 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: BRCA1 (omil)

(viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 17
- (B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTCTCAGA TAACTGGGCC	60
CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA	120
TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAAATGCT ATGCAGAAAA	180
TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC	240
ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT	300
GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC	360
AACTTGTTGA AGAGCTATTG AAAATCATT GTGCTTTTCA GCTTGACACA GGTTTGGAGT	420
ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG	480
AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG	540
AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGA	600
CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG	660
AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG	720
ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG	780
CAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC	840
CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT	900
ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA	960
GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA	1020
AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT	1080
GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG	1140
ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT	1200
CAGAGAATCC TAGAGATACT GAAGATGTTT CTTGGATAAC ACTAAATAGC AGCATTTCAGA	1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG	1320
GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG	1380
AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA	1440
TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT	1500
TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC	1560
TAATTATAGG AGCATTGTGT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA	1620
AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG	1680
CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACAAAC CAAACGGAGC	1740
AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT	1800
CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA	1860
AAACGAAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACCTGAA TTAAATATCC	1920
ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC	1980
ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA	2040
TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA	2100
GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA	2160
GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA	2220
AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT	2280
TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT	2340
CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG	2400
AAAGATCTGT AGAGAGTAGC AGTATTTTAC TGGTACCTGG TACTGATTAT GGCACTCAGG	2460
AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGTAAGGC AAAAACAGAA CCAAATAAAT	2520
GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG	2580
ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC	2640
GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT	2700
TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG	2760
AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT	2820
TTGAATGTGA ACAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC	2880
AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA	2940
ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA	3000
ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAACCCA TATCGTATAC	3060

CACCACTTTT	TCCCATCAAG	TCATTTGTTA	AACTAAATG	TAAGAAAAAT	CTGCTAGAGG	3120
AAAACTTTGA	GGAACATTCA	ATGTCACCTG	AAAGAGAAAT	GGGAAATGAG	AACATTCCAA	3180
GTACAGTGAG	CACAATTAGC	CGTAATAACA	TTAGAGAAAA	TGTTTTTAAA	GGAGCCAGCT	3240
CAAGCAATAT	TAATGAAGTA	GGTTCAGTA	CTAATGAAGT	GGGCTCCAGT	ATTAATGAAA	3300
TAGGTTCCAG	TGATGAAAAC	ATTCAAGCAG	AACTAGGTAG	AAACAGAGGG	CCAAAATTGA	3360
ATGCTATGCT	TAGATTAGGG	GTTTTGCAAC	CTGAGGTCTA	TAAACAAAGT	CTTCCTGGAA	3420
GTAATTGTAA	GCATCCTGAA	ATAAAAAAGC	AAGAATATGA	AGAAGTAGTT	CAGACTGTTA	3480
ATACAGATTT	CTCTCCATAT	CTGATTTTCA	ATAACTTAGA	ACAGCCTATG	GGAAGTAGTC	3540
ATGCATCTCA	GGTTTGTCT	GAGACACCTG	ATGACCTGTT	AGATGATGGT	GAAATAAAGG	3600
AAGATACTAG	TTTTGCTGAA	AATGACATTA	AGGAAAGTTC	TGCTGTTTTT	AGCAAAAGCG	3660
TCCAGAGAGG	AGAGCTTAGC	AGGAGTCCTA	GCCCTTTCAC	CCATACACAT	TTGGCTCAGG	3720
GTTACCGAAG	AGGGGCCAAG	AAATTAGAGT	CCTCAGAAGA	GAAGTTATCT	AGTGAGGATG	3780
AAGAGCTTCC	CTGCTTCCAA	CACTTGTTAT	TTGGTAAAGT	AAACAATATA	CCTTCTCAGT	3840
CTACTAGGCA	TAGCACCGTT	GCTACCGAGT	GTCTGTCTAA	GAACACAGAG	GAGAATTTAT	3900
TATCATTGAA	GAATAGCTTA	AATGACTGCA	GTAACCAGGT	AATATTGGCA	AAGGCATCTC	3960
AGGAACATCA	CCTTAGTGAG	GAAACAAAAT	GTTCTGCTAG	CTTGTTTTCT	TCACAGTGCA	4020
GTGAATTGGA	AGACTTGACT	GCAAATACAA	ACACCCAGGA	TCCTTTCTTG	ATTGGTTCTT	4080
CCAAACAAAT	GAGGCATCAG	TCTGAAAGCC	AGGGAGTTGG	TCTGAGTGAC	AAGGAATTGG	4140
TTTCAGATGA	TGAAGAAAGA	GGAACGGGCT	TGGAAGAAAA	TAATCAAGAA	GAGCAAAGCA	4200
TGGATTCAAA	CTTAGGTGAA	GCAGCATCTG	GGTGTGAGAG	TGAAACAAGC	GTCTCTGAAG	4260
ACTGCTCAGG	GCTATCCTCT	CAGAGTGACA	TTTTAACCAC	TCAGCAGAGG	GATACCATGC	4320
AACATAACCT	GATAAAGCTC	CAGCAGGAAA	TGGCTGAACT	AGAAGCTGTG	TTAGAACAGC	4380
ATGGGAGCCA	GCCTTCTAAC	AGCTACCCTT	CCATCATAAG	TGACTCCTCT	GCCCTTGAGG	4440
ACCTGCGAAA	TCCAGAACAA	AGCACATCAG	AAAAAGCAGT	ATTAAGTTCA	CAGAAAAGTA	4500
GTGAATACCC	TATAAGCCAG	AATCCAGAAG	GCCTTTCTGC	TGACAAGTTT	GAGGTGTCTG	4560
CAGATAGTTC	TACCAGTAAA	AATAAAGAAC	CAGGAGTGGA	AAGGTCATCC	CCTTCTAAAT	4620
GCCCATCATT	AGATGATAGG	TGGTACATGC	ACAGTTGCTC	TGGGAGTCTT	CAGAATAGAA	4680
ACTACCCATC	TCAAGAGGAG	CTCATTAAAG	TTGTTGATGT	GGAGGAGCAA	CAGCTGGAAG	4740
AGTCTGGGCC	ACACGATTTG	ACGGAAACAT	CTTACTTGCC	AAGGCAAGAT	CTAGAGGGAA	4800
CCCCTTACCT	GGAATCTGGA	ATCAGCCTCT	TCTCTGATGA	CCCTGAATCT	GATCCTTCTG	4860
AAGACAGAGC	CCCAGAGTCA	GCTCGTGTTG	GCAACATACC	ATCTTCAACC	TCTGCATTGA	4920

```

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG      4980
ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG      5040
CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG      5100
AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA      5160
CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC      5220
TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTT TGGGTGACCC      5280
AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG      5340
TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT      5400
TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCACA GATCAACTGG      5460
AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG      5520
GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT      5580
TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA      5640
GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA      5700
GCCACTACTG A                                                              5711

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1863 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: BRCA1 (omil)

(viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 17
- (B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1           5           10  o           15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
          20           25           30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
          35           40           45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys

```

50	55	60
Lys Asn Asp Ile Thr 65	Lys Arg Ser Leu Gln 70	Glu Ser Thr Arg Phe Ser 75 80
Gln Leu Val Glu 85	Glu Leu Leu Lys Ile 90	Ile Cys Ala Phe Gln Leu Asp 95
Thr Gly Leu Glu Tyr 100	Ala Asn Ser Tyr 105	Asn Phe Ala Lys Lys Glu Asn 110
Asn Ser Pro Glu His 115	Leu Lys Asp Glu Val 120	Ser Ile Ile Gln Ser Met 125
Gly Tyr Arg Asn Arg 130	Ala Lys Arg Leu Leu 135	Gln Ser Glu Pro Glu Asn 140
Pro Ser Leu Gln Glu 145	Thr Ser Leu Ser Val 150	Gln Leu Ser Asn Leu Gly 155 160
Thr Val Arg Thr 165	Leu Arg Thr Lys Gln 170	Arg Ile Gln Pro Gln Lys Thr 175
Ser Val Tyr Ile 180	Glu Leu Gly Ser Asp 185	Ser Ser Glu Asp Thr Val Asn 190
Lys Ala Thr Tyr Cys 195	Ser Val Gly Asp Gln 200	Glu Leu Leu Gln Ile Thr 205
Pro Gln Gly Thr Arg 210	Asp Glu Ile Ser Leu 215	Asp Ser Ala Lys Lys Ala 220
Ala Cys Glu Phe Ser 225	Glu Thr Asp Val Thr 230	Asn Thr Glu His His Gln 235 240
Pro Ser Asn Asn Asp 245	Leu Asn Thr Thr Glu 250	Lys Arg Ala Ala Glu Arg 255
His Pro Glu Lys Tyr 260	Gln Gly Ser Ser Val 265	Ser Asn Leu His Val Glu 270
Pro Cys Gly Thr Asn 275	Thr His Ala Ser 280	Ser Ser Leu Gln His Glu Asn Ser 285
Ser Leu Leu Leu Thr 290	Lys Asp Arg Met Asn 295	Val Glu Lys Ala Glu Phe 300
Cys Asn Lys Ser Lys 305	Gln Pro Gly Leu Ala 310	Arg Ser Gln His Asn Arg 315 320
Trp Ala Gly Ser Lys 325	Glu Thr Cys Asn Asp 330	Arg Arg Thr Pro Ser Thr 335
Glu Lys Lys Val Asp 340	Leu Asn Ala Asp 345	Pro Leu Cys Glu Arg Lys Glu 350
Trp Asn Lys Gln Lys 355	Leu Pro Cys Ser Glu 360	Asn Pro Arg Asp Thr Glu 365
Asp Val Pro Trp Ile 370	Thr Leu Asn Ser Ser 375	Ile Gln Lys Val Asn Glu 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu
405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg
485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590

Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu

51

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
1045 1050 1055

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Arg Gly
1170 1175 1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185 1190 1195 1200

Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
1205 1210 1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly
1220 1225 1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
1235 1240 1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
1250 1255 1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265 1270 1275 1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
1285 1290 1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
1300 1305 1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
1315 1320 1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
1330 1335 1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345 1350 1355 1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
1365 1370 1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
1380 1385 1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
1395 1400 1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
1410 1415 1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425 1430 1435 1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
1445 1450 1455

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500

Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp

1700	1705	1710
Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met		
1715	1720	1725
Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg		
1730	1735	1740
Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile		
1745	1750	1755
		1760
Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro		
	1765	1770
		1775
Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val		
	1780	1785
		1790
Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val		
	1795	1800
		1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile		
1810	1815	1820
Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp		
1825	1830	1835
		1840
Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro		
	1845	1850
		1855
Gln Ile Pro His Ser His Tyr		
	1860	

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5711 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: BRCA1 (omi2)

(viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 17
- (B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTCTCAGA TAACTGGGCC	60
CCTGCCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA	120
TGGATTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA	180

TCTTAGAGTG	TCCCATCTGT	CTGGAGTTGA	TCAAGGAACC	TGTCTCCACA	AAGTGTGACC	240
ACATATTTTG	CAAATTTTGC	ATGCTGAAAC	TTCTCAACCA	GAAGAAAGGG	CCTTCACAGT	300
GTCCTTTTATG	TAAGAATGAT	ATAACCAAAA	GGAGCCTACA	AGAAAGTACG	AGATTTAGTC	360
AAC TTGTTGA	AGAGCTATTG	AAAATCATT	GTGCTTTTCA	GCTTGACACA	GGTTTGGAGT	420
ATGCAAACAG	CTATAATTTT	GCAAAAAAGG	AAAATAACTC	TCCTGAACAT	CTAAAAAGATG	480
AAGTTTCTAT	CATCCAAAGT	ATGGGCTACA	GAAACCGTGC	CAAAAGACTT	CTACAGAGTG	540
AACCCGAAAA	TCCTTCCTTG	CAGGAAACCA	GTCTCAGTGT	CCAACTCTCT	AACCTTGGA	600
CTGTGAGAAC	TCTGAGGACA	AAGCAGCGGA	TACAACCTCA	AAAGACGTCT	GTCTACATTG	660
AATTGGGATC	TGATTCTTCT	GAAGATACCG	TTAATAAGGC	AACTTATTGC	AGTGTGGGAG	720
ATCAAGAATT	GTTACAAATC	ACCCCTCAAG	GAACCAGGGA	TGAAATCAGT	TTGGATTCTG	780
CAAAAAAGGC	TGCTTG TGAA	TTTTCTGAGA	CGGATGTAAC	AAATACTGAA	CATCATCAAC	840
CCAGTAATAA	TGATTTGAAC	ACCACTGAGA	AGCGTGCAGC	TGAGAGGCAT	CCAGAAAAAGT	900
ATCAGGGTAG	TTCTGTTTCA	AACTTG CATG	TGGAGCCATG	TGGCACAAAT	ACTCATGCCA	960
GCTCATTACA	GCATGAGAAC	AGCAGTTTAT	TACTCACTAA	AGACAGAATG	AATGTAGAAA	1020
AGGCTGAATT	CTGTAATAAA	AGCAAACAGC	CTGGCTTAGC	AAGGAGCCAA	CATAACAGAT	1080
GGGCTGGAAG	TAAGGAAACA	TGTAATGATA	GGCGGACTCC	CAGCACAGAA	AAAAAGGTAG	1140
ATCTGAATGC	TGATCCCCTG	TGTGAGAGAA	AAGAATGGAA	TAAGCAGAAA	CTGCCATGCT	1200
CAGAGAATCC	TAGAGATACT	GAAGATGTTT	CTTGATAAC	ACTAAATAGC	AGCATT CAGA	1260
AAGTTAATGA	GTGGTTTTCC	AGAAGTGATG	AACTGTTAGG	TTCTGATGAC	TCACATGATG	1320
GGGAGTCTGA	ATCAAATGCC	AAAGTAGCTG	ATGTATTGGA	CGTTCTAAAT	GAGGTAGATG	1380
AATATTCTGG	TTCTTCAGAG	AAAATAGACT	TACTG CC CCAG	TGATCCTCAT	GAGGCTTTAA	1440
TATGTAAAAG	TGAAAGAGTT	CACTCCAAAT	CAGTAGAGAG	TAATATTGAA	GACAAAATAT	1500
TTGGGAAAAC	CTATCGGAAG	AAGGCAAGCC	TCCCCAACTT	AAGCCATGTA	ACTGAAAAATC	1560
TAATTATAGG	AGCATTTGTT	ACTGAGCCAC	AGATAATACA	AGAGCGTCCC	CTCACAAATA	1620
AATTAAAGCG	TAAAAGGAGA	CCTACATCAG	GCCTTCATCC	TGAGGATTTT	ATCAAGAAAG	1680
CAGATTTGGC	AGTTCAAAAG	ACTCCTGAAA	TGATAAATCA	GGGAACTAAC	CAAACGGAGC	1740
AGAATGGTCA	AGTGATGAAT	ATTACTAATA	GTGGTCATGA	GAATAAAACA	AAAGGTGATT	1800
CTATT CAGAA	TGAGAAAAAT	CCTAACCCAA	TAGAATCACT	CGAAAAAGAA	TCTGCTTTCA	1860
AAACGAAAGC	TGAACCTATA	AGCAGCAGTA	TAAGCAATAT	GGAAC TCGAA	TTAAATATCC	1920
ACAATTCAAA	AGCACCTAAA	AAGAATAGGC	TGAGGAGGAA	GTCTTCTACC	AGGCATATTC	1980

5371.34.US01

ATGCGCTTGA	ACTAGTAGTC	AGTAGAAATC	TAAGCCCACC	TAATTGTACT	GAATTGCAAA	2040
TTGATAGTTG	TTCTAGCAGT	GAAGAGATAA	AGAAAAAATA	GTACAACCAA	ATGCCAGTCA	2100
GGCACAGCAG	AAACCTACAA	CTCATGGAAG	GTAAAGAACC	TGCAACTGGA	GCCAAGAAGA	2160
GTAACAAGCC	AAATGAACAG	ACAAGTAAAA	GACATGACAG	TGATACTTTC	CCAGAGCTGA	2220
AGTTAACAAA	TGCACCTGGT	TCTTTTACTA	AGTGTTCAAA	TACCAAGTAA	CTTAAAGAAT	2280
TTGTCAATCC	TAGCCTTCCA	AGAGAAGAAA	AAGAAGAGAA	ACTAGAAACA	GTTAAAGTGT	2340
CTAATAATGC	TGAAGACCCC	AAAGATCTCA	TGTTAAGTGG	AGAAAGGGTT	TTGCAAACTG	2400
AAAGATCTGT	AGAGAGTAGC	AGTATTTTAC	TGGTACCTGG	TACTGATTAT	GGCACTCAGG	2460
AAAGTATCTC	GTTACTGGAA	GTTAGCACTC	TAGGGAAGGC	AAAAACAGAA	CCAAATAAAT	2520
GTGTGAGTCA	GTGTGCAGCA	TTTGAAAACC	CCAAGGGACT	AATTCATGGT	TGTTCCAAAG	2580
ATAATAGAAA	TGACACAGAA	GGCTTTAAGT	ATCCATTGGG	ACATGAAGTT	AACCACAGTC	2640
GGGAAACAAG	CATAGAAATG	GAAGAAAGTG	AACTTGATGC	TCAGTATTTG	CAGAATACAT	2700
TCAAGGTTTC	AAAGCGCCAG	TCATTTGCTC	TGTTTTCAAA	TCCAGGAAAT	GCAGAAGAGG	2760
AATGTGCAAC	ATTCTCTGCC	CACTCTGGGT	CCTTAAAGAA	ACAAAGTCCA	AAAGTCACTT	2820
TTGAATGTGA	ACAAAAGGAA	GAAAATCAAG	GAAAGAATGA	GTCTAATATC	AAGCCTGTAC	2880
AGACAGTTAA	TATCACTGCA	GGCTTTCTCT	TGGTTGGTCA	GAAAGATAAG	CCAGTTGATA	2940
ATGCCAAATG	TAGTATCAAA	GGAGGCTCTA	GGTTTTGTCT	ATCATCTCAG	TTCAGAGGCA	3000
ACGAAACTGG	ACTCATTACT	CCAAATAAAC	ATGGACTTTT	ACAAAACCCA	TATCGTATAC	3060
CACCACTTTT	TCCCATCAAG	TCATTTGTGA	AACTAAATG	TAAGAAAAAT	CTGCTAGAGG	3120
AAAACCTTGA	GGAACATTCA	ATGTCACCTG	AAAGAGAAAT	GGGAAATGAG	AACATTCCAA	3180
GTACAGTGAG	CACAATTAGC	CGTAATAACA	TTAGAGAAAA	TGTTTTTAAA	GGAGCCAGCT	3240
CAAGCAATAT	TAATGAAGTA	GGTTCCAGTA	CTAATGAAGT	GGGCTCCAGT	ATTAATGAAA	3300
TAGGTTCCAG	TGATGAAAAC	ATTCAAGCAG	AACTAGGTAG	AAACAGAGGG	CCAAAATTGA	3360
ATGCTATGCT	TAGATTAGGG	GTTTTGCAAC	CTGAGGTCTA	TAAACAAAGT	CTTCCTGGAA	3420
GTAATTGTAA	GCATCCTGAA	ATAAAAAAGC	AAGAATATGA	AGAAGTAGTT	CAGACTGTTA	3480
ATACAGATTT	CTCTCCATAT	CTGATTTTCA	ATAACTTAGA	ACAGCCTATG	GGAAGTAGTC	3540
ATGCATCTCA	GGTTTGTCT	GAGACACCTG	ATGACCTGTT	AGATGATGGT	GAAATAAAGG	3600
AAGATACTAG	TTTTCTTGAA	AATGACATTA	AGGAAAGTTC	TGCTGTTTTT	AGCAAAAGCG	3660
TCCAGAGAGG	AGAGCTTAGC	AGGAGTCCTA	GCCCTTTCAC	CCATACACAT	TTGGCTCAGG	3720
GTTACCGAAG	AGGGGCCAAG	AAATTAGAGT	CCTCAGAAGA	GAACTTATCT	AGTGAGGATG	3780
AAGAGCTTCC	CTGCTTCCAA	CACTTGTTAT	TTGGTAAAGT	AAACAATATA	CCTTCTCAGT	3840

CTACTAGGCA	TAGCACCGTT	GCTACCGAGT	GTCTGTCTAA	GAACACAGAG	GAGAATTTAT	3900
TATCATTGAA	GAATAGCTTA	AATGACTGCA	GTAACCAGGT	AATATTGGCA	AAGGCATCTC	3960
AGGAACATCA	CCTTAGTGAG	GAAACAAAAT	GTTCTGCTAG	CTTGTTTTCT	TCACAGTGCA	4020
GTGAATTGGA	AGACTTGACT	GCAAATACAA	ACACCCAGGA	TCCTTTCTTG	ATTGGTTCTT	4080
CCAAACAAAT	GAGGCATCAG	TCTGAAAGCC	AGGGAGTTGG	TCTGAGTGAC	AAGGAATTGG	4140
TTTCAGATGA	TGAAGAAAGA	GGAACGGGCT	TGGAAGAAAA	TAATCAAGAA	GAGCAAAGCA	4200
TGGATTCAAA	CTTAGGTGAA	GCAGCATCTG	GGTGTGAGAG	TGAAACAAGC	GTCTCTGAAG	4260
ACTGCTCAGG	GCTATCCTCT	CAGAGTGACA	TTTTAACCAC	TCAGCAGAGG	GATACCATGC	4320
AACATAACCT	GATAAAGCTC	CAGCAGGAAA	TGGCTGAACT	AGAAGCTGTG	TTAGAACAGC	4380
ATGGGAGCCA	GCCTTCTAAC	AGCTACCCTT	CCATCATAAG	TGACTCTTCT	GCCCTTGAGG	4440
ACCTGCGAAA	TCCAGAACAA	AGCACATCAG	AAAAAGCAGT	ATTAACCTCA	CAGAAAAGTA	4500
GTGAATACCC	TATAAGCCAG	AATCCAGAAG	GCCTTTCTGC	TGACAAGTTT	GAGGTGTCTG	4560
CAGATAGTTC	TACCAGTAAA	AATAAAGAAC	CAGGAGTGGA	AAGGTCATCC	CCTTCTAAAT	4620
GCCCATCATT	AGATGATAGG	TGGTACATGC	ACAGTTGCTC	TGGGAGTCTT	CAGAATAGAA	4680
ACTACCCATC	TCAAGAGGAG	CTCATTAAGG	TTGTTGATGT	GGAGGAGCAA	CAGCTGGAAG	4740
AGTCTGGGCC	ACACGATTTG	ACGGAAACAT	CTTACTTGCC	AAGGCAAGAT	CTAGAGGGAA	4800
CCCCTTACCT	GGAATCTGGA	ATCAGCCTCT	TCTCTGATGA	CCCTGAATCT	GATCCTTCTG	4860
AAGACAGAGC	CCCAGAGTCA	GCTCGTGTTG	GCAACATACC	ATCTTCAACC	TCTGCATTGA	4920
AAGTTCCCCA	ATTGAAAGTT	GCAGAATCTG	CCCAGGGTCC	AGCTGCTGCT	CATACTACTG	4980
ATACTGCTGG	GTATAATGCA	ATGGAAGAAA	GTGTGAGCAG	GGAGAAGCCA	GAATTGACAG	5040
CTTCAACAGA	AAGGGTCAAC	AAAAGAATGT	CCATGCTGGT	GTCTGGCCTG	ACCCCAGAAG	5100
AATTTATGCT	CGTGTACAAG	TTTGCCAGAA	AACACCACAT	CACTTTAACT	AATCTAATTA	5160
CTGAAGAGAC	TACTCATGTT	GTTATGAAAA	CAGATGCTGA	GTTTGTGTGT	GAACGGACAC	5220
TGAAATATTT	TCTAGGAATT	GCGGGAGGAA	AATGGGTAGT	TAGCTATTTT	TGGGTGACCC	5280
AGTCTATTAA	AGAAAGAAAA	ATGCTGAATG	AGCATGATTT	TGAAGTCAGA	GGAGATGTGG	5340
TCAATGGAAG	AAACCACCAA	GGTCCAAAGC	GAGCAAGAGA	ATCCCAGGAC	AGAAAGATCT	5400
TCAGGGGGCT	AGAAATCTGT	TGCTATGGGC	CCTTCACCAA	CATGCCCACA	GATCAACTGG	5460
AATGGATGGT	ACAGCTGTGT	GGTGCTTCTG	TGGTGAAGGA	GCTTTCATCA	TTCACCCTTG	5520
GCACAGGTGT	CCACCCAATT	GTGGTTGTGC	AGCCAGATGC	CTGGACAGAG	GACAATGGCT	5580
TCCATGCAAT	TGGGCAGATG	TGTGAGGCAC	CTGTGGTGAC	CCGAGAGTGG	GTGTTGGACA	5640

GTGTAGCACT CTACCACTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1863 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: BRCA1 (omi2)

(viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 17
- (B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asp	Leu	Ser	Ala	Leu	Arg	Val	Glu	Glu	Val	Gln	Asn	Val	Ile	Asn	1	5	10	15
Ala	Met	Gln	Lys	Ile	Leu	Glu	Cys	Pro	Ile	Cys	Leu	Glu	Leu	Ile	Lys	20	25	30	
Glu	Pro	Val	Ser	Thr	Lys	Cys	Asp	His	Ile	Phe	Cys	Lys	Phe	Cys	Met	35	40	45	
Leu	Lys	Leu	Leu	Asn	Gln	Lys	Lys	Gly	Pro	Ser	Gln	Cys	Pro	Leu	Cys	50	55	60	
Lys	Asn	Asp	Ile	Thr	Lys	Arg	Ser	Leu	Gln	Glu	Ser	Thr	Arg	Phe	Ser	65	70	75	80
Gln	Leu	Val	Glu	Glu	Leu	Leu	Lys	Ile	Ile	Cys	Ala	Phe	Gln	Leu	Asp	85	90	95	
Thr	Gly	Leu	Glu	Tyr	Ala	Asn	Ser	Tyr	Asn	Phe	Ala	Lys	Lys	Glu	Asn	100	105	110	
Asn	Ser	Pro	Glu	His	Leu	Lys	Asp	Glu	Val	Ser	Ile	Ile	Gln	Ser	Met	115	120	125	
Gly	Tyr	Arg	Asn	Arg	Ala	Lys	Arg	Leu	Leu	Gln	Ser	Glu	Pro	Glu	Asn	130	135	140	
Pro	Ser	Leu	Gln	Glu	Thr	Ser	Leu	Ser	Val	Gln	Leu	Ser	Asn	Leu	Gly	145	150	155	160
Thr	Val	Arg	Thr	Leu	Arg	Thr	Lys	Gln	Arg	Ile	Gln	Pro	Gln	Lys	Thr	165	170	175	
Ser	Val	Tyr	Ile	Glu	Leu	Gly	Ser	Asp	Ser	Ser	Glu	Asp	Thr	Val	Asn	180	185	190	

[illegible]

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590

Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
835 840 845

Glu	Ser	Glu	Leu	Asp	Ala	Gln	Tyr	Leu	Gln	Asn	Thr	Phe	Lys	Val	Ser	850	855	860	
Lys	Arg	Gln	Ser	Phe	Ala	Leu	Phe	Ser	Asn	Pro	Gly	Asn	Ala	Glu	Glu	865	870	875	880
Glu	Cys	Ala	Thr	Phe	Ser	Ala	His	Ser	Gly	Ser	Leu	Lys	Lys	Gln	Ser	885	890	895	
Pro	Lys	Val	Thr	Phe	Glu	Cys	Glu	Gln	Lys	Glu	Glu	Asn	Gln	Gly	Lys	900	905	910	
Asn	Glu	Ser	Asn	Ile	Lys	Pro	Val	Gln	Thr	Val	Asn	Ile	Thr	Ala	Gly	915	920	925	
Phe	Pro	Val	Val	Gly	Gln	Lys	Asp	Lys	Pro	Val	Asp	Asn	Ala	Lys	Cys	930	935	940	
Ser	Ile	Lys	Gly	Gly	Ser	Arg	Phe	Cys	Leu	Ser	Ser	Gln	Phe	Arg	Gly	945	950	955	960
Asn	Glu	Thr	Gly	Leu	Ile	Thr	Pro	Asn	Lys	His	Gly	Leu	Leu	Gln	Asn	965	970	975	
Pro	Tyr	Arg	Ile	Pro	Pro	Leu	Phe	Pro	Ile	Lys	Ser	Phe	Val	Lys	Thr	980	985	990	
Lys	Cys	Lys	Lys	Asn	Leu	Leu	Glu	Glu	Asn	Phe	Glu	Glu	His	Ser	Met	995	1000	1005	
Ser	Pro	Glu	Arg	Glu	Met	Gly	Asn	Glu	Asn	Ile	Pro	Ser	Thr	Val	Ser	1010	1015	1020	
Thr	Ile	Ser	Arg	Asn	Asn	Ile	Arg	Glu	Asn	Val	Phe	Lys	Gly	Ala	Ser	1025	1030	1035	1040
Ser	Ser	Asn	Ile	Asn	Glu	Val	Gly	Ser	Ser	Thr	Asn	Glu	Val	Gly	Ser	1045	1050	1055	
Ser	Ile	Asn	Glu	Ile	Gly	Ser	Ser	Asp	Glu	Asn	Ile	Gln	Ala	Glu	Leu	1060	1065	1070	
Gly	Arg	Asn	Arg	Gly	Pro	Lys	Leu	Asn	Ala	Met	Leu	Arg	Leu	Gly	Val	1075	1080	1085	
Leu	Gln	Pro	Glu	Val	Tyr	Lys	Gln	Ser	Leu	Pro	Gly	Ser	Asn	Cys	Lys	1090	1095	1100	
His	Pro	Glu	Ile	Lys	Lys	Gln	Glu	Tyr	Glu	Glu	Val	Val	Gln	Thr	Val	1105	1110	1115	1120
Asn	Thr	Asp	Phe	Ser	Pro	Tyr	Leu	Ile	Ser	Asp	Asn	Leu	Glu	Gln	Pro	1125	1130	1135	
Met	Gly	Ser	Ser	His	Ala	Ser	Gln	Val	Cys	Ser	Glu	Thr	Pro	Asp	Asp	1140	1145	1150	
Leu	Leu	Asp	Asp	Gly	Glu	Ile	Lys	Glu	Asp	Thr	Ser	Phe	Ala	Glu	Asn	1155	1160	1165	
Asp	Ile	Lys	Glu	Ser	Ser	Ala	Val	Phe	Ser	Lys	Ser	Val	Gln	Arg	Gly	1170	1175	1180	

/

Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg		
1505	1510	1515 1520
Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu		
	1525	1530 1535
Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr		
	1540	1545 1550
Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile		
	1555	1560 1565
Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala		
	1570	1575 1580
Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu		
	1585	1590 1595 1600
Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala		
	1605	1610 1615
Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val		
	1620	1625 1630
Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys		
	1635	1640 1645
Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu		
	1650	1655 1660
Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile		
	1665	1670 1675 1680
Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val		
	1685	1690 1695
Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp		
	1700	1705 1710
Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met		
	1715	1720 1725
Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg		
	1730	1735 1740
Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile		
	1745	1750 1755 1760
Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro		
	1765	1770 1775
Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val		
	1780	1785 1790
Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val		
	1795	1800 1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile		
	1810	1815 1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825 1830 1835 1840
Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
1845 1850 1855
Gln Ile Pro His Ser His Tyr
1860

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5711 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (B) STRAIN: BRCA1 (omi3)

(viii) POSITION IN GENOME:

- (A) CHROMOSOME/SEGMENT: 17
- (B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTCTCAGA TAACTGGGCC	60
CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA	120
TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA	180
TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC	240
ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT	300
GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC	360
AACTTGTTGA AGAGCTATTG AAAATCATTG GTGCTTTTCA GCTTGACACA GGTGGGAGT	420
ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG	480
AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG	540
AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGA	600
CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG	660
AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG	720
ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG	780
CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC	840
CCAGTAATAA TGATTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT	900

ATCAGGGTAG	TTCTGTTTCA	AACTTGCATG	TGGAGCCATG	TGGCACAAAT	ACTCATGCCA	960
GCTCATTACA	GCATGAGAAC	AGCAGTTTAT	TACTCACTAA	AGACAGAATG	AATGTAGAAA	1020
AGGCTGAATT	CTGTAATAAA	AGCAAACAGC	CTGGCTTAGC	AAGGAGCCAA	CATAACAGAT	1080
GGGCTGGAAG	TAAGGAAACA	TGTAATGATA	GGCGGACTCC	CAGCACAGAA	AAAAAGGTAG	1140
ATCTGAATGC	TGATCCCCTG	TGTGAGAGAA	AAGAATGGAA	TAAGCAGAAA	CTGCCATGCT	1200
CAGAGAATCC	TAGAGATACT	GAAGATGTTT	CTTGGATAAC	ACTAAATAGC	AGCATTGAGA	1260
AAGTTAATGA	GTGGTTTTTC	AGAAGTGATG	AACTGTTAGG	TTCTGATGAC	TCACATGATG	1320
GGGAGTCTGA	ATCAAATGCC	AAAGTAGCTG	ATGTATTGGA	CGTTCTAAAT	GAGGTAGATG	1380
AATATTCTGG	TTCTTCAGAG	AAAATAGACT	TACTGGCCAG	TGATCCTCAT	GAGGCTTTAA	1440
TATGTAAAAG	TGAAAGAGTT	CACTCCAAAT	CAGTAGAGAG	TAATATTGAA	GACAAAATAT	1500
TTGGGAAAAC	CTATCGGAAG	AAGGCAAGCC	TCCCCAACTT	AAGCCATGTA	ACTGAAAATC	1560
TAATTATAGG	AGCATTGTGT	ACTGAGCCAC	AGATAATACA	AGAGCGTCCC	CTCACAAATA	1620
AATTAAAGCG	TAAAAGGAGA	CCTACATCAG	GCCTTCATCC	TGAGGATTTT	ATCAAGAAAG	1680
CAGATTTGGC	AGTTCAAAAG	ACTCCTGAAA	TGATAAATCA	GGGAACAAAC	CAAACGGAGC	1740
AGAATGGTCA	AGTGATGAAT	ATTACTAATA	GTGGTCATGA	GAATAAAACA	AAAGGTGATT	1800
CTATTGAGAA	TGAGAAAAAT	CCTAACCCAA	TAGAATCACT	CGAAAAAGAA	TCTGCTTTCA	1860
AAACGAAAGC	TGAACCTATA	AGCAGCAGTA	TAAGCAATAT	GGAACCTGAA	TTAAATATCC	1920
ACAATTCAAA	AGCACCTAAA	AAGAATAGGC	TGAGGAGGAA	GTCTTCTACC	AGGCATATTC	1980
ATGCGCTTGA	ACTAGTAGTC	AGTAGAAATC	TAAGCCCACC	TAATTGTACT	GAATTGCAAA	2040
TTGATAGTTG	TTCTAGCAGT	GAAGAGATAA	AGAAAAAATA	GTACAACCAA	ATGCCAGTCA	2100
GGCACAGCAG	AAACCTACAA	CTCATGGAAG	GTAAAGAACC	TGCAACTGGA	GCCAAGAAGA	2160
GTAACAAGCC	AAATGAACAG	ACAAGTAAAA	GACATGACAG	CGATACTTTC	CCAGAGCTGA	2220
AGTTAACAAA	TGCACCTGGT	TCTTTTACTA	AGTGTTCAAA	TACCAGTGAA	CTTAAAGAAT	2280
TTGTCAATCC	TAGCCTTCCA	AGAGAAGAAA	AAGAAGAGAA	ACTAGAAACA	GTTAAAGTGT	2340
CTAATAATGC	TGAAGACCCC	AAAGATCTCA	TGTTAAGTGG	AGAAAGGGTT	TTGCAAACCTG	2400
AAAGATCTGT	AGAGAGTAGC	AGTATTTTAT	TGGTACCTGG	TACTGATTAT	GGCACTCAGG	2460
AAAGTATCTC	GTTACTGGAA	GTTAGCACTC	TAGGGAAGGC	AAAAACAGAA	CCAAATAAAT	2520
GTGTGAGTCA	GTGTGCAGCA	TTTGAAAACC	CCAAGGGACT	AATTCATGGT	TGTTCCAAAG	2580
ATAATAGAAA	TGACACAGAA	GGCTTTAAGT	ATCCATTGGG	ACATGAAGTT	AACCACAGTC	2640
GGGAAACAAG	CATAGAAATG	GAAGAAAAGT	AACTTGATGC	TCAGTATTTG	CAGAATACAT	2700

TCAAGGTTTC	AAAGCGCCAG	TCATTTGCTC	TGTTTTCAAA	TCCAGGAAAT	GCAGAAGAGG	2760
AATGTGCAAC	ATTCTCTGCC	CACTCTGGGT	CCTTAAAGAA	ACAAAGTCCA	AAAGTCACTT	2820
TTGAATGTGA	ACAAAAGGAA	GAAAATCAAG	GAAAGAATGA	GTCTAATATC	AAGCCTGTAC	2880
AGACAGTTAA	TATCACTGCA	GGCTTTCCTG	TGGTTGGTCA	GAAAGATAAG	CCAGTTGATA	2940
ATGCCAAATG	TAGTATCAAA	GGAGGCTCTA	GGTTTTGTCT	ATCATCTCAG	TTCAGAGGCA	3000
ACGAAACTGG	ACTCATTACT	CCAAATAAAC	ATGGACTTTT	ACAAAACCCA	TATCGTATAC	3060
CACCACTTTT	TCCCATCAAG	TCATTTGTTA	AAACTAAATG	TAAGAAAAAT	CTGCTAGAGG	3120
AAAACCTTTGA	GGAACATTCA	ATGTCACCTG	AAAGAGAAAT	GGGAAATGAG	AACATTCCAA	3180
GTACAGTGAG	CACAATTAGC	CGTAATAACA	TTAGAGAAAA	TGTTTTTAAA	GAAGCCAGCT	3240
CAAGCAATAT	TAATGAAGTA	GGTTCAGTA	CTAATGAAGT	GGGCTCCAGT	ATTAATGAAA	3300
TAGGTTCCAG	TGATGAAAAC	ATTCAAGCAG	AACTAGGTAG	AAACAGAGGG	CCAAAATTGA	3360
ATGCTATGCT	TAGATTAGGG	GTTTTGCAAC	CTGAGGTCTA	TAAACAAAGT	CTTCCTGGAA	3420
GTAATTGTAA	GCATCCTGAA	ATAAAAAAGC	AAGAATATGA	AGAAGTAGTT	CAGACTGTTA	3480
ATACAGATTT	CTCTCCATAT	CTGATTTTCA	ATAACTTAGA	ACAGCCTATG	GGAAGTAGTC	3540
ATGCATCTCA	GGTTTGTCT	GAGACACCTG	ATGACCTGTT	AGATGATGGT	GAAATAAAGG	3600
AAGATACTAG	TTTTGCTGAA	AATGACATTA	AGGAAAGTTC	TGCTGTTTTT	AGCAAAAGCG	3660
TCCAGAAAGG	AGAGCTTAGC	AGGAGTCCTA	GCCCTTTCAC	CCATACACAT	TTGGCTCAGG	3720
GTTACCGAAG	AGGGGCCAAG	AAATTAGAGT	CCTCAGAAGA	GAACCTATCT	AGTGAGGATG	3780
AAGAGCTTCC	CTGCTTCCAA	CACCTGTAT	TTGGTAAAGT	AAACAATATA	CCTTCTCAGT	3840
CTACTAGGCA	TAGCACCGTT	GCTACCGAGT	GTCTGTCTAA	GAACACAGAG	GAGAATTTAT	3900
TATCATTGAA	GAATAGCTTA	AATGACTGCA	GTAACGAGGT	AATATTGGCA	AAGGCATCTC	3960
AGGAACATCA	CCTTAGTGAG	GAAACAAAAT	GTTCTGCTAG	CTTGTTTTCT	TCACAGTGCA	4020
GTGAATTGGA	AGACTTGACT	GCAAATACAA	ACACCCAGGA	TCCTTTCTTG	ATTGGTTCTT	4080
CCAAACAAAT	GAGGCATCAG	TCTGAAAGCC	AGGGAGTTGG	TCTGAGTGAC	AAGGAATTGG	4140
TTTCAGATGA	TGAAGAAAGA	GGAACGGGCT	TGGAAGAAAA	TAATCAAGAA	GAGCAAAGCA	4200
TGGATTCAAA	CTTAGGTGAA	GCAGCATCTG	GGTGTGAGAG	TGAAACAAGC	GTCTCTGAAG	4260
ACTGCTCAGG	GCTATCCTCT	CAGAGTGACA	TTTTAACCAC	TCAGCAGAGG	GATACCATGC	4320
AACATAACCT	GATAAAGCTC	CAGCAGGAAA	TGGCTGAACT	AGAAGCTGTG	TTAGAACAGC	4380
ATGGGAGCCA	GCCTTCTAAC	AGCTACCCTT	CCATCATAAG	TGACTCTTCT	GCCCTTGAGG	4440
ACCTGCGAAA	TCCAGAACAA	AGCACATCAG	AAAAAGCAGT	ATTAACCTCA	CAGAAAAGTA	4500
GTGAATACCC	TATAAGCCAG	AATCCAGAAG	GCCTTTCTGC	TGACAAGTTT	GAGGTGTCTG	4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT	4620
GGCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA	4680
ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG	4740
AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA	4800
CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG	4860
AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA	4920
AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGAGTCC AGCTGCTGCT CATACTACTG	4980
ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG	5040
CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG	5100
AATTTATGCT CGTGTAACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA	5160
CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC	5220
TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTT TGGGTGACCC	5280
AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG	5340
TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT	5400
TCAGGGGGCT AGAAATCTGT TGCTATGGG CTTTACCAA CATGCCACA GATCAACTGG	5460
AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG	5520
GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT	5580
TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA	5640
GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCAG ATCCCCACA	5700
GCCACTACTG A	5711

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1863 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: not relevant
- (ii) MOLECULE TYPE: protein
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
 - (B) STRAIN: BRCA1 (omi3)
- (viii) POSITION IN GENOME:
 - (A) CHROMOSOME/SEGMENT: 17
 - (B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Asp	Leu	Ser	Ala	Leu	Arg	Val	Glu	Glu	Val	Gln	Asn	Val	Ile	Asn	
1					5					10					15	
Ala	Met	Gln	Lys	Ile	Leu	Glu	Cys	Pro	Ile	Cys	Leu	Glu	Leu	Ile	Lys	
		20					25					30				
Glu	Pro	Val	Ser	Thr	Lys	Cys	Asp	His	Ile	Phe	Cys	Lys	Phe	Cys	Met	
	35					40					45					
Leu	Lys	Leu	Leu	Asn	Gln	Lys	Lys	Gly	Pro	Ser	Gln	Cys	Pro	Leu	Cys	
	50					55					60					
Lys	Asn	Asp	Ile	Thr	Lys	Arg	Ser	Leu	Gln	Glu	Ser	Thr	Arg	Phe	Ser	
65					70					75					80	
Gln	Leu	Val	Glu	Glu	Leu	Leu	Lys	Ile	Ile	Cys	Ala	Phe	Gln	Leu	Asp	
			85					90						95		
Thr	Gly	Leu	Glu	Tyr	Ala	Asn	Ser	Tyr	Asn	Phe	Ala	Lys	Lys	Glu	Asn	
		100						105						110		
Asn	Ser	Pro	Glu	His	Leu	Lys	Asp	Glu	Val	Ser	Ile	Ile	Gln	Ser	Met	
		115					120						125			
Gly	Tyr	Arg	Asn	Arg	Ala	Lys	Arg	Leu	Leu	Gln	Ser	Glu	Pro	Glu	Asn	
	130					135					140					
Pro	Ser	Leu	Gln	Glu	Thr	Ser	Leu	Ser	Val	Gln	Leu	Ser	Asn	Leu	Gly	
145					150					155					160	
Thr	Val	Arg	Thr	Leu	Arg	Thr	Lys	Gln	Arg	Ile	Gln	Pro	Gln	Lys	Thr	
			165					170						175		
Ser	Val	Tyr	Ile	Glu	Leu	Gly	Ser	Asp	Ser	Ser	Glu	Asp	Thr	Val	Asn	
		180						185					190			
Lys	Ala	Thr	Tyr	Cys	Ser	Val	Gly	Asp	Gln	Glu	Leu	Leu	Gln	Ile	Thr	
		195					200					205				
Pro	Gln	Gly	Thr	Arg	Asp	Glu	Ile	Ser	Leu	Asp	Ser	Ala	Lys	Lys	Ala	
	210					215					220					
Ala	Cys	Glu	Phe	Ser	Glu	Thr	Asp	Val	Thr	Asn	Thr	Glu	His	His	Gln	
225					230					235					240	
Pro	Ser	Asn	Asn	Asp	Leu	Asn	Thr	Thr	Glu	Lys	Arg	Ala	Ala	Glu	Arg	
			245						250					255		
His	Pro	Glu	Lys	Tyr	Gln	Gly	Ser	Ser	Val	Ser	Asn	Leu	His	Val	Glu	
		260						265					270			
Pro	Cys	Gly	Thr	Asn	Thr	His	Ala	Ser	Ser	Leu	Gln	His	Glu	Asn	Ser	
		275					280					285				
Ser	Leu	Leu	Leu	Thr	Lys	Asp	Arg	Met	Asn	Val	Glu	Lys	Ala	Glu	Phe	
	290					295					300					
Cys	Asn	Lys	Ser	Lys	Gln	Pro	Gly	Leu	Ala	Arg	Ser	Gln	His	Asn	Arg	
305					310					315					320	

Trp	Ala	Gly	Ser	Lys	Glu	Thr	Cys	Asn	Asp	Arg	Arg	Thr	Pro	Ser	Thr	325	330	335	
Glu	Lys	Lys	Val	Asp	Leu	Asn	Ala	Asp	Pro	Leu	Cys	Glu	Arg	Lys	Glu	340	345	350	
Trp	Asn	Lys	Gln	Lys	Leu	Pro	Cys	Ser	Glu	Asn	Pro	Arg	Asp	Thr	Glu	355	360	365	
Asp	Val	Pro	Trp	Ile	Thr	Leu	Asn	Ser	Ser	Ile	Gln	Lys	Val	Asn	Glu	370	375	380	
Trp	Phe	Ser	Arg	Ser	Asp	Glu	Leu	Leu	Gly	Ser	Asp	Asp	Ser	His	Asp	385	390	395	400
Gly	Glu	Ser	Glu	Ser	Asn	Ala	Lys	Val	Ala	Asp	Val	Leu	Asp	Val	Leu	405	410	415	
Asn	Glu	Val	Asp	Glu	Tyr	Ser	Gly	Ser	Ser	Glu	Lys	Ile	Asp	Leu	Leu	420	425	430	
Ala	Ser	Asp	Pro	His	Glu	Ala	Leu	Ile	Cys	Lys	Ser	Glu	Arg	Val	His	435	440	445	
Ser	Lys	Ser	Val	Glu	Ser	Asn	Ile	Glu	Asp	Lys	Ile	Phe	Gly	Lys	Thr	450	455	460	
Tyr	Arg	Lys	Lys	Ala	Ser	Leu	Pro	Asn	Leu	Ser	His	Val	Thr	Glu	Asn	465	470	475	480
Leu	Ile	Ile	Gly	Ala	Phe	Val	Thr	Glu	Pro	Gln	Ile	Ile	Gln	Glu	Arg	485	490	495	
Pro	Leu	Thr	Asn	Lys	Leu	Lys	Arg	Lys	Arg	Arg	Pro	Thr	Ser	Gly	Leu	500	505	510	
His	Pro	Glu	Asp	Phe	Ile	Lys	Lys	Ala	Asp	Leu	Ala	Val	Gln	Lys	Thr	515	520	525	
Pro	Glu	Met	Ile	Asn	Gln	Gly	Thr	Asn	Gln	Thr	Glu	Gln	Asn	Gly	Gln	530	535	540	
Val	Met	Asn	Ile	Thr	Asn	Ser	Gly	His	Glu	Asn	Lys	Thr	Lys	Gly	Asp	545	550	555	560
Ser	Ile	Gln	Asn	Glu	Lys	Asn	Pro	Asn	Pro	Ile	Glu	Ser	Leu	Glu	Lys	565	570	575	
Glu	Ser	Ala	Phe	Lys	Thr	Lys	Ala	Glu	Pro	Ile	Ser	Ser	Ser	Ile	Ser	580	585	590	
Asn	Met	Glu	Leu	Glu	Leu	Asn	Ile	His	Asn	Ser	Lys	Ala	Pro	Lys	Lys	595	600	605	
Asn	Arg	Leu	Arg	Arg	Lys	Ser	Ser	Thr	Arg	His	Ile	His	Ala	Leu	Glu	610	615	620	
Leu	Val	Val	Ser	Arg	Asn	Leu	Ser	Pro	Pro	Asn	Cys	Thr	Glu	Leu	Gln	625	630	635	640
Ile	Asp	Ser	Cys	Ser	Ser	Ser	Glu	Glu	Ile	Lys	Lys	Lys	Lys	Tyr	Asn				

645	650	655
Gln Met Pro Val Arg His Ser Arg 660	Asn Leu Gln Leu Met 665	Glu Gly Lys 670
Glu Pro Ala Thr Gly Ala Lys Lys 675	Ser Asn Lys Pro Asn 680	Gln Thr 685
Ser Lys Arg His Asp Ser 690	Asp Thr Phe Pro Glu 695	Lys Leu Thr Asn 700
Ala Pro Gly Ser Phe Thr Lys Cys 705	Ser Asn Thr Ser Glu 710	Leu Lys Glu 715
Phe Val Asn Pro Ser Leu Pro Arg 725	Glu Glu Lys Glu Glu 730	Lys Leu Glu 735
Thr Val Lys Val Ser Asn Asn Ala 740	Glu Asp Pro Lys Asp 745	Leu Met Leu 750
Ser Gly Glu Arg Val Leu Gln Thr 755	Glu Arg Ser Val Glu 760	Ser Ser Ser 765
Ile Ser Leu Val Pro Gly Thr Asp 770	Tyr Gly Thr Gln Glu 775	Ser Ile Ser 780
Leu Leu Glu Val Ser Thr Leu Gly 785	Lys Ala Lys Thr Glu 790	Pro Asn Lys 795
Cys Val Ser Gln Cys Ala Ala Phe 805	Glu Asn Pro Lys Gly 810	Leu Ile His 815
Gly Cys Ser Lys Asp Asn Arg Asn 820	Asp Thr Glu Gly Phe 825	Lys Tyr Pro 830
Leu Gly His Glu Val Asn His Ser 835	Arg Glu Thr Ser Ile 840	Glu Met Glu 845
Glu Ser Glu Leu Asp Ala Gln Tyr 850	Leu Gln Asn Thr Phe 855	Lys Val Ser 860
Lys Arg Gln Ser Phe Ala Leu Phe 865	Ser Asn Pro Gly Asn 870	Ala Glu Glu 875
Glu Cys Ala Thr Phe Ser Ala His 885	Ser Gly Ser Leu Lys 890	Lys Gln Ser 895
Pro Lys Val Thr Phe Glu Cys Glu 900	Gln Lys Glu Glu Asn 905	Gln Gly Lys 910
Asn Glu Ser Asn Ile Lys Pro Val 915	Gln Thr Val Asn Ile 920	Thr Ala Gly 925
Phe Pro Val Val Gly Gln Lys Asp 930	Lys Pro Val Asp Asn 935	Ala Lys Cys 940
Ser Ile Lys Gly Gly Ser Arg Phe 945	Cys Leu Ser Ser Gln 950	Phe Arg Gly 955
Asn Glu Thr Gly Leu Ile Thr Pro 965	Asn Lys His Gly Leu 970	Leu Gln Asn 975

[illegible]

1300	1305	1310
Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser 1315 1320 1325		
Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp 1330 1335 1340		
Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser 1345 1350 1355 1360		
Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr 1365 1370 1375		
Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu 1380 1385 1390		
Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln 1395 1400 1405		
Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln 1410 1415 1420		
Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu 1425 1430 1435 1440		
Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr 1445 1450 1455		
Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu 1460 1465 1470		
Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn 1475 1480 1485		
Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu 1490 1495 1500		
Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg 1505 1510 1515 1520		
Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu 1525 1530 1535		
Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr 1540 1545 1550		
Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile 1555 1560 1565		
Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala 1570 1575 1580		
Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu 1585 1590 1595 1600		
Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Ser Pro Ala Ala 1605 1610 1615		
Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val 1620 1625 1630		

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
1700 1705 1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
1715 1720 1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
1730 1735 1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745 1750 1755 1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
1765 1770 1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
1780 1785 1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
1795 1800 1805

Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
1810 1815 1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825 1830 1835 1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
1845 1850 1855

Gln Ile Pro His Ser His Tyr
1860

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGTTGTCA TTTTATAAAC CTTT

24

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 2R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TGTCCTTTTCT TCCCTAGTAT GT

22

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 3F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCCTGACACA GCAGACATTT A

21

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 3R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGGATTTTC GTTCTCACTT A

21

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 5F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTCTTAAGGG CAGTTGTGAG

20

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 5R-M13* primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCTACTGT GGTGCTTCC

20

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 6/7F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTTATTTTAG TGTCCTTAAA AGG

23

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 6R

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTTCATGGAC AGCACTTGAG TG

22

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 7F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CACAACAAAG AGCATACATA GGG

23

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 6/7R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGGGTTTAC TCTGTAGAAG

20

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 8F1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCTCTTCAG GAGGAAAAGC A

21

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 8R1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCTGCCTACC ACAAATACAA A

21

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 9F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCACAGTAGA TGCTCAGTAA ATA

23

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 (B) STRAIN: 9R primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TAGGAAAATA CCAGCTTCAT AGA

23

- (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: not relevant
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 (B) STRAIN: 10F primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGGTCAGCTT TCTGTAATCG

20

- (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: not relevant
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 (B) STRAIN: 10R primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GATCTACCC ACTCTCTTCT TCAG

24

- (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: not relevant
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 (B) STRAIN: 11AF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCACCTCCAA GGTGTATCA

19

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11AR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGTTATGTTG GCTCCTTGCT

20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11BF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CACTAAAGAC AGAATGAATC TA

22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11BR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GAAGAACCAG AATATTCATC TA

22

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 11CF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGATGGGGAG TCTGAATCAA

20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 11CR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TCTGCTTTCT TGATAAAATC CT

22

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 11DF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AGCGTCCCCCT CACAAATAAA

20

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 11DR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TCAAGCGCAT GAATATGCCT

20

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 11EF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTATAAGCAA TATGGAACTC GA

22

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 11ER primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TTAAGTTCAC TGGTATTGAA CA

23

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
(B) STRAIN: 11FF primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GACAGCGATA CTTTCCCAGA

20

- (2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

- (C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
(B) STRAIN: 11FR primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TGGAACAACC ATGAATTAGT C

21

- (2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
(B) STRAIN: 11GF primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGAAGTTAGC ACTCTAGGGA

20

- (2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:

(B) STRAIN: 11GR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GCAGTGATAT TAACTGTCTG TA

22

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11HF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TGGGTCCTTA AAGAAACAAA GT

22.

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11HR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TCAGGTGACA TTGAATCTTC C

21

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11IF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCACTTTTTTC CCATCAAGTC A

21

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 11IR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TCAGGATGCT TACAATTACT TC

22

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 11JF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CAAAATTGAA TGCTATGCTT AGA

23

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 11JR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCGGTAACCC TGAGCCAAAT

20

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 11KF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCAAAAGCGT CCAGAAAGGA

20

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 11KR-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TATTTGCAGT CAAGTCTTCC AA

22

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 11LF-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTAATATTGG CAAAGGCATC T

21

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11LR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TAAAATGTGC TCCCCAAAAG CA

22

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 12F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTCCTGCCAA TGAGAAGAAA

20

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 12R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGTCAGCAAA CCTAAGAATG T

21

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AATGGAAAGC TTCTCAAAGT A

21

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATGTTGGAGC TAGGTCCTTA C

21

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 14F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CTAACCTGAA TTATCACTAT CA

22

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 14R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTGTATAAAT GCCTGTATGC A

21

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 15F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TGGCTGCCCA GGAAGTATG

19

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 15R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AACCAGAATA TCTTTATGTA GGA

23

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (vi) ORIGINAL SOURCE:
 - (B) STRAIN: 16F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

AATTCTTAAC AGAGACCAGA AC

22

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 16R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AAAACTCTTT CCAGAATGTT GT

22

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 17F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GTGTAGAACG TGCAGGATTG

20

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 17R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCGCCTCATG TGGTTTTA

18

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 18F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GGCTCTTTAG CTTCTTAGGA C

21

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 18R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GAGACCATTT TCCCAGCATC

20

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 19F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CTGTCATTCT TCCTGTGCTC

20

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 19R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CATTGTTAAG GAAAGTGGTG C

21

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 20F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

ATATGACGTG TCTGCTCCAC

20

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 20R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GGGAATCCAA ATTACACAGC

20

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 21F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

AAGCTCTTCC TTTTGAAG TC

22

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 21R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GTAGAGAAAT AGAATAGCCT CT

22

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 22F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

TCCCATTGAG AGGTCTTGCT

20

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 22R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GAGAAGACTT CTGAGGCTAC

20

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 23F-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TGAAGTGACA GTTCCAGTAG T

21

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 23R-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CATTTTAGCC ATTCAATCAA CAA

23

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (B) STRAIN: 24F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ATGAATTGAC ACTAATCTCT GC

22

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 24R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GTAGCCAGGA CAGTAGAAGG A

21

GTAGCCAGGA CAGTAGAAGG A